Assignment 3 64060

Nate Cvelbar

2023-10-16

```
#install.packages("caret")
library(caret)

#install.packages("ISLR") # only install if needed
library(ISLR)

#install.packages("e1071") # only install if needed
library(e1071)

#install.packages("pivottabler") # only install if needed
library(pivottabler)

#install.packages("MASS")
#install.packages("reshape2")
#install.packages("reshape")

library(MASS)
library(reshape2)
library(reshape)
```

```
#Assignment 3
#Nate Cvelbar
#BA-64060

#File taken online from course Assignment 3
#Loading the dataset
UB=read.csv('C:/Users/Owner/Documents/UniversalBank.csv')
#Remove unwanted columns
UB<-UB[,-1:-9]
UB<-UB[,-2:-3]
set.seed(111)

#Separate into train and test
Index_Train<-createDataPartition(UB$Personal.Loan, p=0.6, list=FALSE)
Train <-UB[Index_Train,]
Test  <-UB[-Index_Train,]
```

```
#Create pivot table for training data
MSet=melt(Train,c("CreditCard","Personal.Loan"),variable= "Online")
CSet=cast(MSet,CreditCard+Personal.Loan~Online)
CSet
```

```
##   CreditCard Personal.Loan Online
## 1          0             0   1896
## 2          0             1    205
## 3          1             0    812
## 4          1             1     87
```

```
#Total number of customers with CC, Online, and Loan is 143, out of 5000 total.
#This means that the probably of accepting a Loan having a CC and Online is  0.0286, or 2.86%
```

```
#Create 2 pivot tables for the training data, for Loan as function of Online and for Loan as function o
MSetL=melt(Train,c("Personal.Loan"),variable = "Online")
MSetC=melt(Train,c("CreditCard"),variable = "Online")
CSetL=cast(MSetL,Personal.Loan~Online)
CSetC=cast(MSetC,CreditCard~Online)

#Remove Unwanted columns
CSetL[,-3]
```

```
##   Personal.Loan Online
## 1             0   2708
## 2             1    292
```

```
CSetC[,-2]
```

```
##   CreditCard Online
## 1          0   2101
## 2          1    899
```

```
#Compute the various questions
#i. P(CC = 1 | Loan = 1)
table(Train[,c(1,3)])
```

```
##              CreditCard
## Personal.Loan    0    1
##             0 1896  812
##             1  205   87
```

```
#P=87/(87+205)=29.8%
```

```
#ii. P(Online = 1 | Loan = 1)
table(Train[,c(1,2)])
```

```
##              Online
## Personal.Loan    0    1
##             0 1112 1596
##             1  112  180
```

```
#P=180/(180+112)=61.6%
```

```r
#iii. P(Loan = 1)
table(Train[,c(1)])
```

```
##
##    0    1
## 2708  292
```

```r
#P=292/(2725+292)=9.7%
```

```r
#iv. P(CC = 1 | Loan = 0)
table(Train[,c(1,3)])
```

```
##               CreditCard
## Personal.Loan    0    1
##             0 1896  812
##             1  205   87
```

```r
#P=812/(812+1896)=30.0%
```

```r
#v. P(Online = 1 | Loan = 0)
table(Train[,c(1,2)])
```

```
##               Online
## Personal.Loan    0    1
##             0 1112 1596
##             1  112  180
```

```r
#P=1596/(1596+1112)=58.9%
```

```r
#vi. P(Loan = 0)
table(Train[,c(1)])
```

```
##
##    0    1
## 2708  292
```

```r
#P=1-P(Loan=1)=100%-9.7%=90.3%
```

```r
#P(Loan = 1 | CC= 1, Online = 1)
#P=[P(CC = 1 | Loan = 1)*P(Online = 1 | Loan = 1)*P(Loan = 1)]/{[P(CC = 1 | Loan = 1)*P(Online = 1 | Lo
#P=10.04%
#Unfortunately, this value is very different from the 2.86% I calculated in step B. The more accurate e
```

```r
#To find this, I would need to run all the the calcations in a way that corresponds to the step D. Then
```